



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

This is to certify that the following application annexed hereto is a true copy from the records of the Korean Intellectual Property Office.

출원 번호 : 10-2002-0082912
Application Number

출원 년 월 일 : 2002년 12월 23일
Date of Application DEC 23, 2002

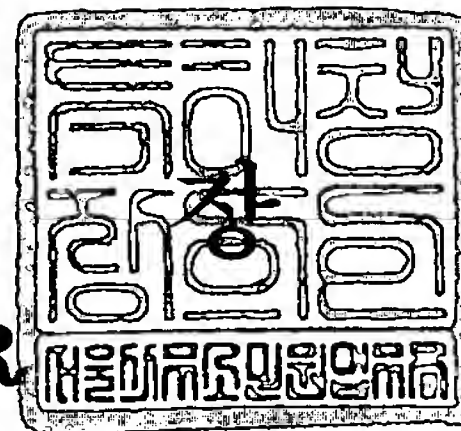
출원인 : 한국전자통신연구원
Applicant(s) Electronics and Telecommunications Research Institute



2003 년 09 월 01 일

특 허 청

COMMISSIONER



【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0011
【제출일자】	2002. 12. 23
【발명의 명칭】	네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치
【발명의 영문명칭】	A network-storage apparatus for high-speed streaming data transmission through network
【출원인】	
【명칭】	한국전자통신연구원
【출원인코드】	3-1998-007763-8
【대리인】	
【성명】	권태복
【대리인코드】	9-2001-000347-1
【포괄위임등록번호】	2001-057650-1
【대리인】	
【성명】	이화익
【대리인코드】	9-1998-000417-9
【포괄위임등록번호】	1999-021997-1
【발명자】	
【성명의 국문표기】	김성운
【성명의 영문표기】	KIM, Seong-Woon
【주민등록번호】	611222-1120412
【우편번호】	305-752
【주소】	대전광역시 유성구 송강동 송강청솔아파트 205동 601호
【국적】	KR
【발명자】	
【성명의 국문표기】	박종원
【성명의 영문표기】	PARK, Chong-Won
【주민등록번호】	580103-1675622

【우편번호】	302-791
【주소】	대전광역시 서구 월평동 누리아파트 111동 1402호
【국적】	KR
【발명자】	
【성명의 국문표기】	정병권
【성명의 영문표기】	JUNG, Byung-Kwon
【주민등록번호】	710320-1711113
【우편번호】	305-350
【주소】	대전광역시 유성구 가정동 236-1
【국적】	KR
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 권태복 (인) 대리인 이화익 (인)
【수수료】	
【기본출원료】	20 면 29,000 원
【가산출원료】	5 면 5,000 원
【우선권주장료】	0 건 0 원
【심사청구료】	9 항 397,000 원
【합계】	431,000 원
【감면사유】	정부출연연구기관
【감면후 수수료】	215,500 원
【기술이전】	
【기술양도】	희망
【실시권 허여】	희망
【기술지도】	희망
【첨부서류】	1. 요약서·명세서(도면)_1통

【요약서】**【요약】**

본 발명은 네트워크를 통해 수신된 데이터를 zero copy 형태로 디스크에 저장하며, 디스크의 스트리밍 데이터를 zero copy 형태로 네트워크를 통해 많은 사용자에게 고속으로 전송하여 인터넷 망을 통해 스트리밍 데이터를 빠르게 송/수신할 수 있도록 하는 네트워크-스토리지 연결 장치에 관한 것이다.

본 발명은 호스트 프로세서와 NSU간의 버스 트랜잭션을 전달하는 주변장치 버스 브릿지와, 외부의 주변장치 버스와 분리되는 내부 주변장치 버스와, 디스크 스토리지에 대한 데이터 읽기 또는 쓰기를 관리하는 디스크 제어기와, 디스크 스토리지와 네트워크간의 전송 데이터가 저장되는 주변메모리와, 주변메모리의 데이터 저장 및 출력을 관리하는 주변메모리 제어기와, 네트워크로 전송될 데이터를 주변 메모리에서 읽어와서 패킷형태로 가공하여 네트워크로 전송하고, 네트워크로부터의 디스크에 저장할 데이터를 주변메모리에 저장하는 TOE로 구성된다.

【대표도】

도 2

【색인어】

NSU, TOE, PCI 버스, TCP/IP, iSCSI, Zero Copy, 호스트 프로세서

【명세서】

【발명의 명칭】

네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치{A network-storage apparatus for high-speed streaming data transmission through network}

【도면의 간단한 설명】

도 1은 본 발명이 적용되는 인터넷 서버 컴퓨터 시스템에 대한 구성도.

도 2는 본 발명에 따른 네트워크-스토리지 연결 장치에 대한 내부 구성도.

도 3은 본 발명에 따른 네트워크 패킷 처리 부분의 수신 프로세스에 대한 흐름도.

도 4는 본 발명에 따른 네트워크 패킷 처리 부분의 송신 프로세서에 대한 흐름도.

도 5는 본 발명에 따른 스토리지 디스크 제어기의 프로세스 흐름도.

도 6은 본 발명에 따른 주변 메모리 장치에 대한 접근 흐름도.

<도면의 주요부분에 대한 부호의 설명>

10: 호스트 프로세서 20: 프로세서 버스

30: 버스 브릿지 40: 메인 메모리

50: 호스트 주변장치 버스 60: 디스크 스토리지

70: 네트워크 장치 100: 네트워크-스토리지 연결장치

110: PCI 브릿지 120: PCI 버스

130: 디스크 제어기 140: PCI 메모리 제어기

150: TOE 170: PCI 메모리

180: MAC

【발명의 상세한 설명】

【발명의 목적】

【발명이 속하는 기술분야 및 그 분야의 종래기술】

<16> 본 발명은 네트워크상의 스트리밍 데이터를 송수신하기 위한 네트워크-스토리지 연결 장치에 관한 것이며, 보다 상세히는 네트워크를 통해 수신된 데이터를 메모리 복사 등을 통하지 않고 직접 사용자에게 전달하거나 디스크에 저장하고, 또한 디스크의 스트리밍 데이터를 메모리 복사 등을 통하지 않고 직접 네트워크를 통해 데이터를 전송함으로써 인터넷 망을 통해 스트리밍 데이터를 빠르게 송/수신할 수 있도록 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치에 관한 것이다.

<17> 최근, 인터넷을 통해서 VoD (Video-on-Demand), NoD (News-on-Demand) 등과 같은 스트리밍 데이터를 전송하고자 하는 서비스들의 사용이 급격히 증가하고 있다.

- <18> 또한, 인터넷 네트워크의 전송 대역이 높아 감에 따라 스트리밍 데이터를 전송해야 할 컴퓨터 서버의 기능은 더욱 높은 성능의 처리 효율을 요구한다. 즉, 서버에 연결된 사용자의 숫자만큼 인터넷 상에서 서버로부터 내려 받는 데이터의 양은 급격히 증가한다.
- <19> 이렇게 미디어 서버가 다루어야 하는 데이터는 급격히 증가하고 있지만 서버의 운영체제는 빠르게 발전하는 하드웨어의 발전 속도를 따라가지 못하고 있는 실정이다.
- <20> 이러한 네트워크 서비스를 원활히 처리하기 위해서는 기존의 운영체제 레벨에서 수행되던 TCP/IP 기능을 외부의 TOE (TCP/IP Offload Engine)에서 수행하고자 하는 경향이 늘어나고 있다.
- <21> 즉 TOE 하드웨어가 운영체제에서 수행하던 TCP/IP 기능을 대신 수행함으로써, 서버의 프로세서를 90% 이상 사용하여 서버 성능의 주요 감소 요인이었던 TCP/IP 수행 부분을 덜게 되어 전체적인 시스템 성능을 향상시킬 수 있다. 또한, 빠른 TOE의 수행으로 네트워크를 통한 사용자의 요구에 빠르게 응답할 수 있게 되었다.
- <22> TOE는 이러한 TCP/IP 기능을 하드웨어적으로 처리할 수 있도록 구성되어 있다. 실제로 TOE를 ASIC으로 구현하여 제품화 한 경우도 있다.
- <23> 이에 대해서는 특허 USA-P 6,427,173 B1호(발명의 명칭; Intelligent network Interfaced Device and System For Accelerated Communication)가 출원된 바 있다.

- <24> 하지만, 이와 같은 TOE를 사용하는 종래 방법도 스트리밍 데이터를 전송하기 위한 디스크 접근에 대한 고려가 없기 때문에 여전히 스트리밍 데이터를 전송함에 있어 호스트 프로세서에 많은 부하가 걸리게 되는 문제점이 있다.
- <25> 또 다른 방법으로서 스토리지를 네트워크 망을 통하여 공유할 수 있는 방법으로 iSCSI를 사용한다. iSCSI는 SCSI 블록 I/O 프로토콜을 일반적인 TCP/IP 프로토콜을 사용하여 네트워크를 통해서 SCSI 블록 I/O 프로토콜을 전송할 수 있도록 해 준다.
- <26> iSCSI에 관련된 특허는 많이 출원된 상태이며, 특허 USA-P 6,400,730호 (Method and apparatus for transferring data between IP network devices and SCSI and fibre channel devices over an IP network)가 출원된 바 있다.
- <27> 하지만, 이와 같은 iSCSI 방식으로 시스템을 구현하여도 서버에 많은 사용자가 연결될 경우에 이를 서비스하는 데는 호스트 프로세서의 부담이 여전히 상태이다.
- <28> 한편, 인터넷을 통하여 대량의 스트리밍 데이터를 전송할 경우 해당 콘텐츠의 복사가 컴퓨터 시스템 내에서 여러 번 일어난다. 이러한 무계획적인 복사는 인터넷 서버의 성능을 감소시킬 뿐 아니라 QoS를 보장할 수 있는 스트리밍을 제공하는 것을 어렵게 하는 문제점이 있다.
- <29> 따라서, 이러한 빠른 네트워크를 이용하는 많은 사용자의 QoS를 만족시켜 줄 수 있는 컴퓨터의 하드웨어의 새로운 구조가 필요하게 되었으며, 이러한 하드

웨어의 구조는 네트워크의 처리 성능을 향상시키고, 스트리밍 데이터의 빠른 접근을 수행할 수 있는 형태가 되어야 한다.

【발명이 이루고자 하는 기술적 과제】

<30> 따라서, 본 발명은 상술한 종래의 문제점을 해결하기 위한 것으로서, 본 발명의 목적은 네트워크를 통해 수신된 데이터를 zero copy 형태로 디스크에 저장하며, 또한 디스크의 스트리밍 데이터를 zero copy 형태로 네트워크를 통해 많은 사용자들에게 고속으로 전송함으로써, 인터넷 망을 통해 스트리밍 데이터를 빠르게 송/수신할 수 있도록 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치를 제공하는데 있다.

<31> 상기 본 발명의 목적을 달성하기 위한 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치는, 인터넷 서버 컴퓨터 시스템의 다수의 디스크 스토리지와 네트워크 장치에 대한 스트리밍 데이터를 처리하는 네트워크-스토리지 연결장치에 있어서, 호스트 프로세서로부터의 버스 트랜잭션을 내부 주변장치 버스로 전달하고, NSU 내부에서 진행되는 호스트 프로세서 또는 메인 메모리에 대한 버스 트랜잭션을 버스 브릿지로 전달하는 주변장치 버스 브릿지; NSU 외부의 주변장치 버스와 분리되고, NSU 내부 각 장치간의 데이터 전송을 위한 내부 주변장치 버스; NSU에 연결되는 다수의 디스크 스토리지를 제어하여 디스크 스토리지에 대한 데이터 읽기 또는 쓰기를 관리하는 디스크 제어기; 디스크 스토리지와 네트워크간의 전송 데이터가 저장되는 주변메모리; 상기 주변메모리를 제어하여 디스크 스토리지와 네트워크간의 전송 데이터를 저장

시키거나 출력시키는 주변메모리 제어기; 및 네트워크로 전송될 데이터를 상기 주변 메모리에서 읽어와서 패킷형태로 가공하여 네트워크로 전송하고, 네트워크로부터 수신되는 데이터를 상기 주변메모리 제어기를 통해 상기 주변메모리에 저장하는 TOE;로 구성된다.

【발명의 구성 및 작용】

- <32> 이하, 본 발명에 따른 실시예를 첨부한 도면을 참조하여 상세히 설명하기로 한다.
- <33> 도 1은 본 발명이 적용되는 인터넷 서버 컴퓨터 시스템에 대한 개략적인 구성도이다.
- <34> 도 1에 도시된 바와 같이, 인터넷 서버 컴퓨터 시스템은 일반적으로 호스트 프로세서(10)들이 프로세서 버스(20)에 연결되어 있으며, 이 프로세서 버스(20)에는 메인 메모리(40)를 접근하거나 또 다른 주변장치 버스(50)를 연결하는 버스 브릿지(30)가 연결되는 구조로 이루어진다.
- <35> 또한, 상기 버스 브릿지(30)는 호스트 프로세서(10)들에 의해 수행되는 명령어에 대한 버스 트랜잭션이 프로세서 버스(20)에 나타나면 이를 분석하여 어느 장치를 접근하는 버스 트랜잭션 인지를 알아낸다.
- <36> 대부분 이러한 과정은 어드레스 버스에서 구동되는 어드레스로써 판단되며, 어드레스 영역이 각 장치별로 할당되어 있다. 만약에 프로세서 버스 트랜잭션이

메모리 영역을 접근하며 메인 메모리(40)를 접근하고 주변장치를 접근하면 주변 장치 버스(50)로 버스 트랜잭션을 넘겨주는 역할을 수행한다.

<37> 많이 사용되는 주변장치 버스(50)로는 PCI가 대표적이다.

<38> 본 발명에 따른 네트워크-스토리지 연결장치(100)는 이러한 주변장치 버스(50)에 연결되어 있으며, 주변장치에 연결 가능한 장치 개수만큼 장착이 가능하다.

<39> 또한, 상기 네트워크-스토리지 연결장치(Network-Storage Unit, 이하에서는 간단히 NSU라 함, 100)는 위로는 주변장치 버스(50)의 인터페이스에 맞게 동작되며, 아래로는 디스크 스토리지(60), 및 이더넷(Ethernet)과 같은 네트워크 장치(70)에 연결된다.

<40> 따라서, 버스 브릿지(30)를 통해서 전달되는 호스트 프로세서(10)의 요구를 받아 들여, 디스크를 접근하는 경우에는 디스크 스토리지(60)를 읽거나 쓰기를 하고, 네트워크의 접근을 요구하는 경우에는 네트워크 장치(70)를 통해서 데이터 패킷을 전송하는 역할을 수행한다.

<41> 한편, 도 2는 상기 설명된 바와 같은 역할을 수행하는 본 발명의 네트워크-스토리지 연결 장치(NSU)에 대한 내부 구성도이다.

<42> 도 2에 도시된 바와 같이, 본 발명에 따른 NSU(100)는 크게 주변장치 버스 브릿지(110)와, 디스크 제어기(130)와, 주변메모리 제어기(140) 및 주변메모리(170)와, TOE(TCP/IP Offload Engine, 150) 및 MAC(Medium Access Control, 180)로 구성된다.

- <43> 상기 주변장치 버스 브릿지(110)는 호스트 프로세서(10)가 요청한 버스 트랜잭션이 NSU(100)를 접근하는 경우에 버스 브릿지(30)에 의해 주변장치 버스(50)를 통해 전달된 이 트랜잭션을 다시 NSU(100) 내부의 주변장치 버스(120)로 전달하는 기능을 수행한다.
- <44> 따라서, 상기 주변장치 버스 브릿지(110)는 내부에 어드레스를 지정하는 레지스터를 가지고 있으며 또한 시스템의 초기화 작업이 진행되는 동안에 NSU(100)에서 사용되는 여러가지 자원에 대한 정보를 저장하고 있다.
- <45> 만약, 주변장치 버스로써 PCI가 사용되는 경우에는 상기 주변장치 버스 브릿지(110)는 PCI 브릿지 역할을 수행하게 된다.
- <46> 이 PCI 브릿지(110)는 장치(NSU) 내부에서 진행되는 버스 트랜잭션에 대하여 이를 호스트 프로세서(10)에게 알리기 위해 사용되거나, 혹은 메인 메모리(40)를 접근하는 트랜잭션이 수행되는 경우에 이를 버스 브릿지(30)에게 전달한다.
- <47> 또한, 프로세서에 의해 진행되는 버스 트랜잭션이 PCI 장치를 접근하는 경우에 이를 받아서 트랜잭션을 전달하는 기능을 수행한다.
- <48> 본 발명의 특징 중 하나는 NSU(100) 내부 주변장치 버스(120)는 외부의 주변장치 버스(50)와 분리되어 구성하는 것이다.
- <49> 이렇게 주변장치버스가 분리되어 구성되는 경우에는 NSU(100) 내부에서 진행되는 네트워크-스토리지 간의 버스 트랜잭션이 호스트 주변장치 버스(50)에 실제로 나타나지 않게 된다.

- <50> 이것은 호스트 주변장치 버스(50)의 버스 트래픽을 급격하게 감소시킬 수 있게 한다.
- <51> 만약, 내부에 주변장치버스(120)가 없다면, 모든 트래픽이 호스트 주변장치 버스(50)로 물리게 되며, 실제로 이 버스(50)가 PCI 버스인 경우라도 이 곳이 병목 지점이 된다. NSU(100)가 여러 개가 연결되는 경우에는 이러한 현상이 더욱 심각하게 되며, 실제로 PCI 버스의 대역폭이 32 비트 버스인 경우에는 대략 133 Mbyte/sec 정도가 되지만, 이것은 산술적으로 네트워크의 1 Gbps 정도 데이터 처리를 겨우 할 수 있는 정도이다.
- <52> 그러나, 본 발명의 경우처럼 NSU(100) 내부에 버스(120)을 두게 되면, 비록 PCI 버스의 대역폭이 네트워크의 대역폭에 따라 가지 못하는 경우에도 실제로 호스트 주변장치 버스(50)로 전달되는 데이터 양을 급격히 줄일 수 있기 때문에 시스템의 성능을 감소시키지 않게 하는 효과가 있다.
- <53> 또한, NSU(100) 내부에 있는 상기 디스크 제어기(130)는 여러 개의 디스크(60)를 병렬로 연결할 수 있도록 되어 있다. 이는 실제로 디스크 접근이 또 다른 병목 현상이 되는 것을 막기 위한 것이다.
- <54> 상기 디스크 제어기(130)는 내부에 디스크(60)를 스트라이핑하는 기능을 포함하도록 되어 있어 여러 개의 디스크(60)에 데이터를 나누어서 저장할 수 있도록 되어 있다.
- <55> 디스크 스트라이핑(striping)은 호스트 프로세서(10)나 네트워크(70)를 통해 디스크(60)로 대량의 데이터를 저장하는 요구가 들어 오는 경우에 대량의 데

이터를 소량의 데이터 군으로 나누어서 조금씩 각각의 디스크에 저장하는 방법이다. 이렇게 저장하게 되면 동시에 모든 디스크를 사용하게 되는 효과가 있다. 이는 한 개의 디스크에 대량의 데이터를 한꺼번에 써넣는 동안에 또 다른 디스크 접근 요구가 생기는 경우에 이를 처리해 줄 수 없는 단점을 보완하는 기능이다.

<56> 대부분 디스크 인터페이스 버스(160)로 많이 사용되는 것이 SCSI 프로토콜이다. SCSI 프로토콜의 데이터 전송 능력은 최근 들어 160 MBps 혹은 320 MBps 정도로 이르고 있다. 이 정도의 데이터 전송 대역폭을 최대한 사용하기 위한 방법으로 디스크 스트라이핑 방식이 사용된다.

<57> 본 발명의 또 다른 중요한 특징은 NSU(100) 내부에 고유의 메모리(170)를 두어 스토리지와 네트워크 간의 데이터 전송을 완충하는 역할을 수행하게 하는 것이다.

<58> 네트워크를 통해 데이터가 전송되어야 하는 경우에는 상기 주변메모리 제어기(140)는 디스크(60)로부터 전송된 데이터를 주변메모리(170)에 저장한다. 저장된 메모리의 내용은 다시 TOE(150)를 통해서 네트워크(70)로 전송이 된다.

<59> 이렇게 NSU(100) 내부에 메모리(170)가 있으면 스토리지와 네트워크 간의 데이터 전송 속도가 일치하지 않는데 생기는 문제점을 제거하는 동시에, 대량의 데이터를 캐싱하는 기능을 제공할 수 있다. 이는 대량의 PCI 메모리(170)가 NSU(100) 내부에 구성되어 있어, 네트워크(70)를 통해 디스크(60) 접근을 요청하는 경우에 또 다시 디스크(60)를 접근하지 않고 바로 메모리(170) 내의 데이터를 바로 제공함으로써 네트워크 전송 효율을 극대화 시킬 수 있다.

- <60> 상기 주변메모리 제어기(140)는 상기 주변메모리(170)를 제어하는 기능을 수행하며 내부에 주변메모리의 크기를 지정하는 레지스터를 내장하고 있다. 또한 대량의 데이터를 DMA 방식으로 전송할 수 있는 기능도 함께 내장하고 있다.
- <61> 또한, 상기 TOE(150)는 일반적인 TOE의 기능을 가지고 있는 동시에 NSU(100)의 구조에 맞는 기능을 가지고 있다.
- <62> 보통의 TOE는 TCP/IP 프로토콜을 하드웨어적으로 분리하여 수행함으로써 네트워크를 통한 패킷 송수신을 빠르게 하는 기능을 가지고 있다.
- <63> 본 발명에서 제안되고 있는 TOE는 일반적인 TOE의 기능 뿐 아니라 또 다른 기능을 포함하고 있다.
- <64> 상기 TOE의 기능을 두 가지 경우에 대하여 각각 설명하면 다음과 같다.
- <65> 먼저, 디스크(60) 내의 컨텐츠 정보를 네트워크(70)를 통해서 패킷을 전송해야 하는 경우에는 반드시 네트워크를 통해 전송 요구를 보내기 전에 보내야 할 데이터를 상기 주변메모리(170)에 먼저 저장해 두어야 한다.
- <66> 디스크(60)에서 읽혀져 주변메모리(170)에 저장된 데이터는 네트워크를 통해 전송을 수행하라는 명령이 전달되면, 주변메모리(170)로부터 읽혀져 와서 네트워크 전송에 필요한 여러 가지 정보를 첨가하여 네트워크(70)를 통해 패킷 형태로 전송된다.
- <67> 이때, 상기 TOE(150)는 메인 메모리(40)를 접근하지 않게 되고 주변 메모리(170)를 접근하게 되어 호스트 주변장치버스(50)로의 트랜잭션 전송이 일

어나지 않게 되므로 시스템 내의 호스트 프로세서(10)의 동작을 전혀 방해하지 않게 되어 전체적으로 시스템의 성능을 향상하게 한다.

<68> 두 번째로, 네트워크(70)를 통해 디스크(60)로 컨텐츠가 저장되어야 하는 경우, 상기 TOE(150)는 이를 직접 디스크(60)로 써넣는 것이 아니라 바로 주변메모리 제어기(140)를 통해서 주변메모리(170)로 데이터를 써넣는다.

<69> 이 방법 역시 호스트 주변장치 버스(50)를 사용하지 않는 효과를 발생시킨다. TOE(150)는 네트워크로부터 전송되어진 패킷 정보를 해독하는 능력을 가지고 있음과 동시에 이 패킷 데이터가 디스크로 저장되어야 하는 지를 알 수 있다. 이러한 기능은 결국 컨텐츠 정보를 바로 디스크(60)로 전송할 수 있는 길을 제공하는 것이 된다. 이렇게 디스크(60)의 정보를 읽어와서 네트워크(70)를 통해 전송하거나, 네트워크(70)를 통해 전송된 데이터를 디스크(60)로 바로 저장할 수 있도록 경로를 제공함으로써 호스트 프로세서(10)에게 인터럽트, 혹은 로드를 주지 않게 된다.

<70> 이러한 네트워크-스토리지 offload 기능은 본 발명에 따른 NSU(100)의 기본 기능이다.

<71> 한편, 도 3은 상기 TOE(150) 내에서 패킷이 네트워크를 통해서 수신되는 경우 그 패킷의 처리 과정을 보여주는 흐름도이다.

<72> 도 3에 도시된 바와 같이, 먼저 수신 패킷이 발생할 것을 호스트 프로세서(10)에서 이미 알고 있기 때문에, 먼저 DSB를 생성하여 TOE(150)에 저장시킨다.(S301)

- <73> DSB(Disk Save Buffer Table)는 수신되는 패킷 중에서 디스크(60)로 바로 전송되어야 할 패킷 데이터들의 정보를 가지고 있는 테이블이다.
- <74> 이렇게 DSB가 먼저 생성된 후에, 네트워크(70)를 통하여 패킷이 수신된다.(S302)
- <75> 패킷이 수신되면, 상기 TOE(150)는 먼저 그 패킷이 정상적인지를 검사한다.(S303)
- <76> 또한, 패킷의 검사가 완료되면 이 패킷이 디스크(60)로 저장이 가능한 것인지를 확인하고,(S304) 만약 디스크 저장이 되지 않는 일반적인 패킷인 경우에는 네트워크 스택으로 해당 패킷 데이터를 전송한다.(S305)
- <77> 만약 디스크 저장이 가능한 경우에는 이미 형성되어 있는 DSB에 디스크 저장을 기다리는 정보가 있는지를 확인한다.(S306)
- <78> 이때, DSB와 일치되는 정보를 가지고 있으면 이를 NSU(100)의 주변메모리(170)로 전송한다.(S307)
- <79> 일치되는 정보를 가지고 있지 않으면, 역시 이를 일반적인 네트워크 스택으로 전송하여 일반적인 패킷 처리 과정을 따르게 한다.(S308,S309)
- <80> 한편, 도 4는 상기 TOE(150)를 통하여 패킷을 전송하는 과정을 보여주는 흐름도이다.
- <81> 일단, 디스크(60)에서 읽혀진 스트리밍 데이터는 주변메모리(170) 내에 저장되어 있는 상태이다. 주변메모리(170)로의 데이터 전송이 완료되면 호스트 프

로세서(10)로부터 네트워크(70)로 스트리밍 데이터를 전송할 것을 요청하는 명령이 내려온다.(S401)

<82> 이 데이터 전송명령이 내려오면, 상기 TOE(150)는 주변메모리(170)에서 해당 데이터를 읽어 온다.(S402)

<83> 읽어 온 데이터는 네트워크를 통해 패킷 상태로 전송하기 위해서 TOE (150) 내에서 패킷 형태로 구성이 된다.

<84> 이렇게 구성된 패킷이 준비가 완료되었는지를 확인하여,(S403) 완료가 덜 되었으면, 주변메모리(170) 내의 데이터를 읽어와 계속 준비를 하여야 하며,(S404) 준비가 완료되면 이를 네트워크(70)를 통하여 전송한다.(S405)

<85> 이렇게 해당되는 데이터의 양만큼 전송이 완료되었는지를 확인하여 전송이 완료되면, 패킷 송신이 끝나게 된다.(S406)

<86> 한편, 도 5는 상기 디스크 제어기(130)를 통하여 디스크로 데이터를 쓰거나 읽기를 하는 과정을 보여주는 흐름도이다.

<87> 호스트 프로세서(10)로부터 디스크 접근에 대한 명령이 있으면,(S501) 상기 디스크 제어기(130)는 먼저 그 명령이 디스크 쓰기인지를 확인한다.(S502)

<88> 디스크 쓰기인 경우에는 주변장치버스(120)를 통해 쓰기 데이터를 수신한다.(S503) 이때, 쓰기 데이터를 스트라이핑하는 단위로 쪼개어 이를 디스크(60)로 쓰기를 수행한다.(S504)

<89> 읽기인 경우에는 디스크의 스트라이핑 단위로 디스크(60)를 읽는다.(S505)

- <90> 이렇게 작은 단위로 여러 개의 디스크를 병렬로 접근하는 방식은 디스크의 접근을 파이프라인 방식으로 접근하게 함으로써, 디스크의 접근 성능을 향상시킨다.
- <91> 따라서 스트라이핑 단위만큼 디스크 접근이 완료되면, (S506) 즉시 다음의 디스크 접근 데이트를 처리할 수 있는 기능을 포함한다.
- <92> 도 6은 상기 메모리 제어기(140)를 통하여 주변메모리(170)를 접근하는 과정을 보여주는 흐름도이다.
- <93> 주변 메모리(170)를 사용하기 위해서 먼저 메모리 제어기(140)의 내부에 어드레스 범위 지정 레지스터를 초기화 시켜야 한다. (S601)
- <94> 즉, NSU(100) 내의 주변메모리(170)의 영역을 지정할 수 있도록 상기 메모리 제어기(140)를 초기화한다. 이때 메모리 접근에 관한 정보를 가지고 있는 메모리 테이블도 같이 정립하여야 한다.
- <95> 메모리 접근에 관한 정보는 현재 주변메모리(170)의 어떤 부분이 프로세서에 의해 사용되고 있는 지를 확인할 수 있는 테이블이다.
- <96> 메모리 제어기(140)에 주변메모리 접근을 위한 명령이 들어오면, 이를 메모리 테이블에 해당 어드레스와 함께 사용하고 있음을 표시해 둔다. (S602)
- <97> 이 정보는 디스크 접근의 요청이 있는 경우에 이미 디스크(60)를 읽어와서 메모리(170)에 저장되어 있는 지를 확인하기 위해서 사용된다. 즉, 주변메모리(170)를 네트워크 캐싱으로 사용할 수 있도록 한 것이다.

- <98> 그리고, 그 접근 명령이 메모리 쓰기인지를 확인하여, (S603) 쓰기이면 수신된 쓰기 데이터를 주변메모리(170)에 써넣고, (S604) 읽기이면 주변메모리(170)를 읽어서 해당 데이터를 메모리 제어기(140)로 넘겨준다. (S605)
- <99> 그리고, 메모리 접근이 완료되었는지를 확인하여, (S606) 완료가 된 상태이면 메모리 테이블의 값을 삭제시킨다. (S607)
- <100> 메모리 테이블의 값이 유용한 상태로 남아 있으면, 거기에 해당되는 메모리 값이 계속 유효한 것으로 인식이 되므로 반드시 메모리 사용이 완료된 경우에는 이를 클리어 시킨다.

【발명의 효과】

- <101> 상술한 바와 같이 본 발명에 따른 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치는, 네트워크 및 디스크의 스트리밍 데이터를 zero copy 형태로 처리함으로써, 서버의 호스트 프로세서의 부담을 줄이고 인터넷 망을 통해 스트리밍 데이터를 빠르게 송/수신할 수 있도록 한다.
- <102> 또한, 디스크의 스트리밍 데이터에 대해 메모리 복사를 줄이고, 프로세서의 간섭을 최소화한 상태로 처리함으로써 고품질의 스트리밍 데이터를 QoS 지원할 수 있도록 한다.
- <103> 이상에서 설명한 것은 본 발명에 따른 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치를 실시하기 위한 하나의

실시예에 불과한 것으로서, 본 발명은 상기한 실시예에 한정되지 않고, 이하의 특허청구의 범위에서 청구하는 본 발명의 요지를 벗어남이 없이 당해 발명이 속하는 분야에서 통상의 지식을 가진 자라면 누구든지 다양한 변경 실시가 가능한 범위까지 본 발명의 기술적 정신이 있다고 할 것이다.

【특허청구범위】**【청구항 1】**

인터넷 서버 컴퓨터 시스템의 다수의 디스크 스토리지와 네트워크 장치에 대한 스트리밍 데이터를 처리하는 네트워크-스토리지 연결장치에 있어서,

호스트 프로세서로부터의 버스 트랜잭션을 내부 주변장치 버스로 전달하고, NSU 내부에서 진행되는 호스트 프로세서 또는 메인 메모리에 대한 버스 트랜잭션을 버스 브릿지로 전달하는 주변장치 버스 브릿지;

NSU 외부의 주변장치 버스와 분리되고, NSU 내부 각 장치간의 데이터 전송을 위한 내부 주변장치 버스;

NSU 에 연결되는 다수의 디스크 스토리지를 제어하여 디스크 스토리지에 대한 데이터 읽기 또는 쓰기를 관리하는 디스크 제어기;

디스크 스토리지와 네트워크간의 전송 데이터가 저장되는 주변메모리;

상기 주변메모리를 제어하여 디스크 스토리지와 네트워크간의 전송 데이터를 저장시키거나 출력시키는 주변메모리 제어기; 및

네트워크로 전송될 데이터를 상기 주변 메모리에서 읽어와서 패킷형태로 가공하여 네트워크로 전송하고, 네트워크로부터 수신되는 데이터를 상기 주변메모리 제어기를 통해 상기 주변메모리에 저장하는 TOE;로 구성되는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 2】

제 1항에 있어서, 상기 주변장치 버스는 PCI 버스가 사용되고, 상기 주변장치 버스 브릿지는 PCI 브릿지 역할을 수행하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치

【청구항 3】

제 1항에 있어서, 상기 디스크 제어기는 디스크 인터페이스 버스를 통해 다수의 디스크 스토리지와 병렬로 연결되고 파이프라인 방식으로 데이터에 접근하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 4】

제 1항에 있어서, 상기 디스크 제어기는 스트라이핑 방식으로 다수의 디스크 스토리지에 대한 데이터 읽기 또는 쓰기를 수행하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 5】

제 1항에 있어서, 상기 주변메모리 제어기는 네트워크 송수신 데이터를 캐싱할 수 있도록 메모리 테이블을 구성하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 6】

제 1항에 있어서, 상기 주변메모리 제어기는 내부에 주변메모리의 크기를 지정하는 레지스터를 내장하고 대량의 데이터를 DMA 방식으로 전송하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 7】

제 1항에 있어서, 상기 주변메모리 제어기는 상기 주변메모리에 대한 접근이 완료되면 그 메모리 테이블 값을 삭제하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 8】

제 1항에 있어서, 상기 TOE는, 네트워크로부터 수신되는 데이터 패킷을 처리함에 있어, 먼저 수신 패킷중 디스크로 바로 전송되어야 할 패킷 데이터에 대한 정보를 갖는 DSB 테이블을 생성한 후, 디스크에 저장 가능한 데이터 패킷이 DSB

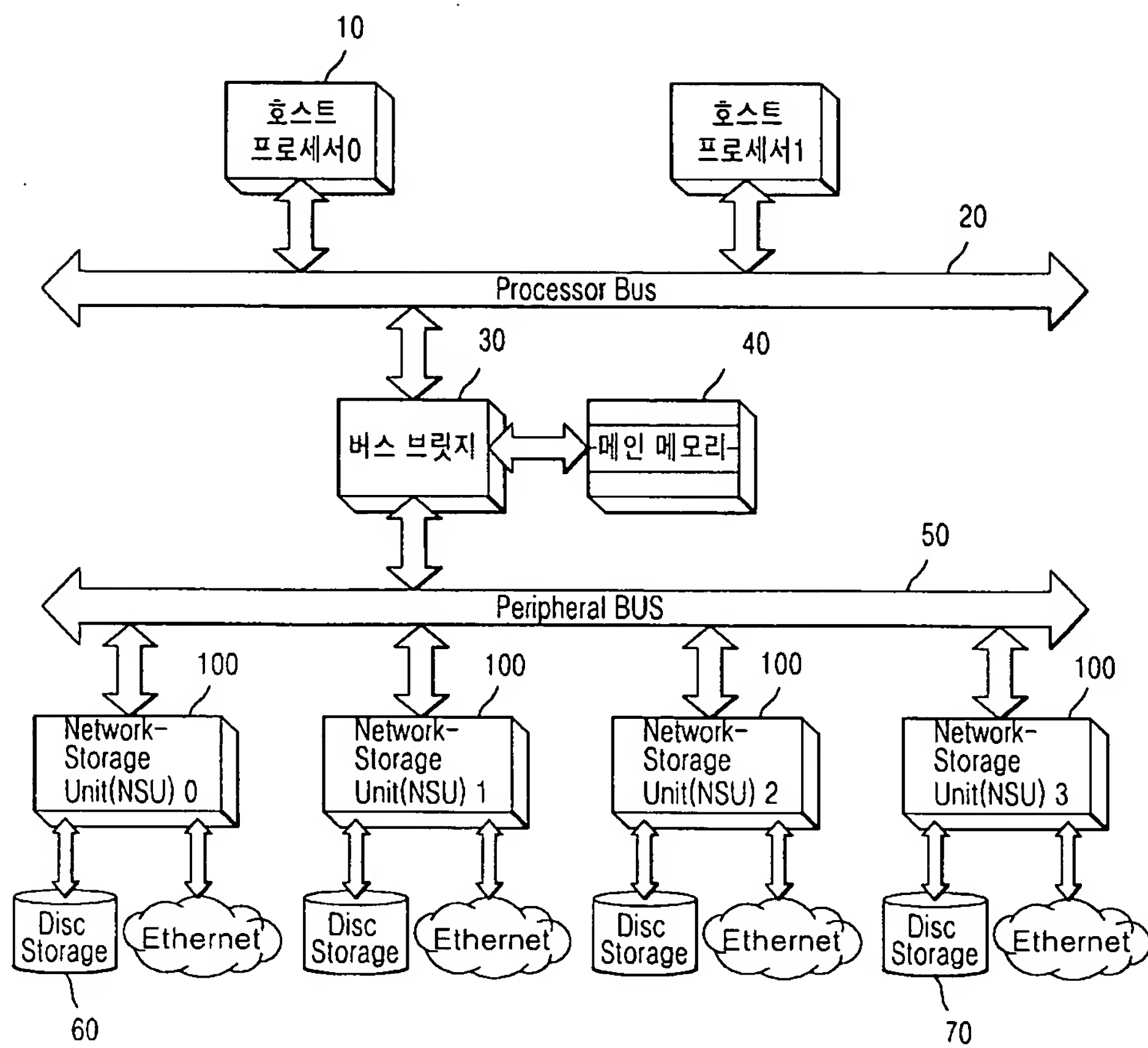
와 일치하는 정보를 가지면 상기 주변 메모리로 전송하여 저장하고, 그 외의 경우는 일반적인 네트워크 스택으로 전송하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【청구항 9】

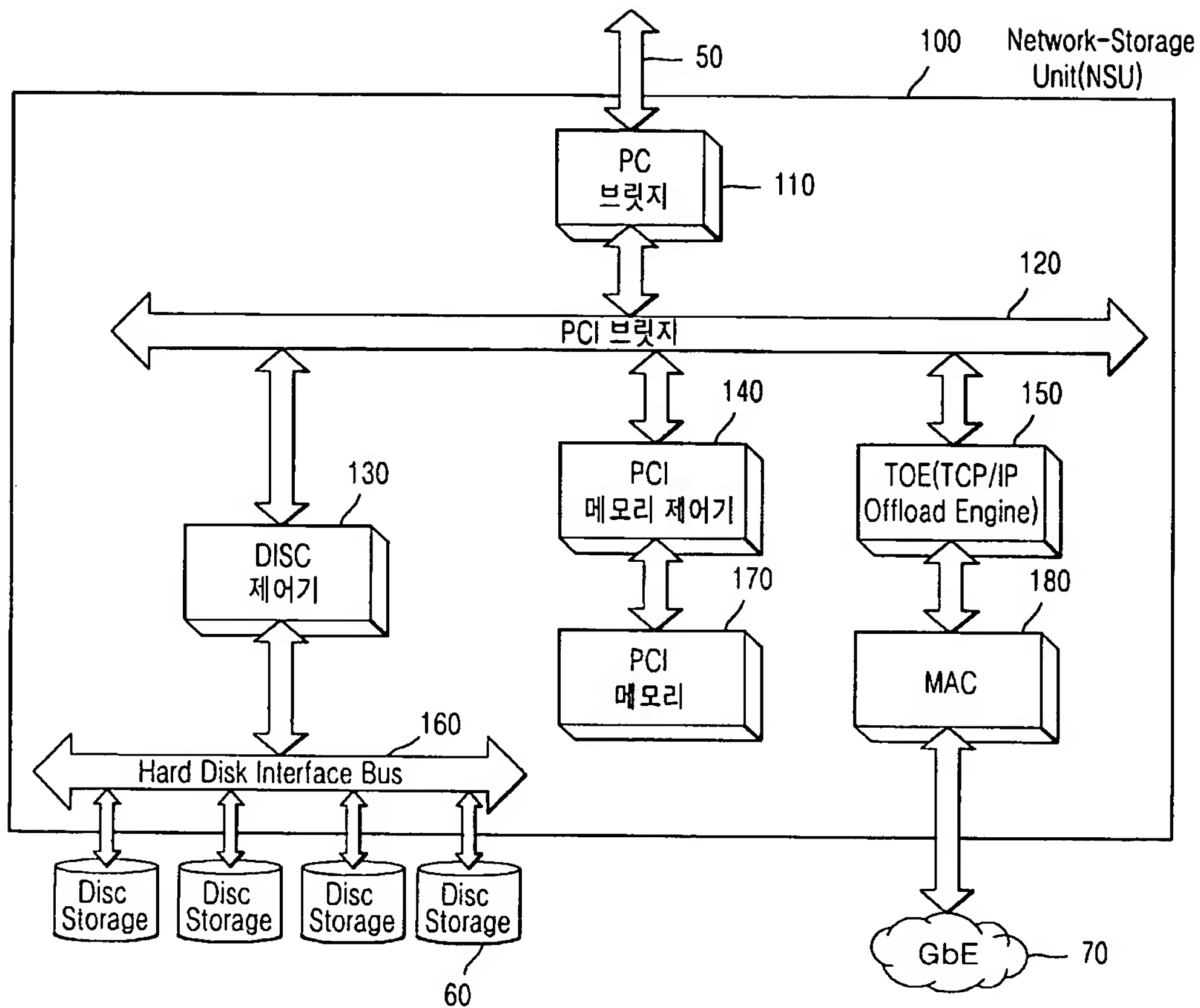
제 1항에 있어서, 상기 TOE는, 네트워크로 데이터 패킷을 전송함에 있어, 전송할 데이터가 상기 주변메모리에 저장된 상태에서 호스트 프로세서로부터 데이터 전송명령이 내려오면, 상기 주변메모리에서 그 데이터를 읽어온 후 이를 패킷 가공하여 네트워크로 전송하는 것을 특징으로 하는 네트워크를 통해 스트리밍 데이터를 고속으로 송수신하기 위한 네트워크-스토리지 연결 장치.

【도면】

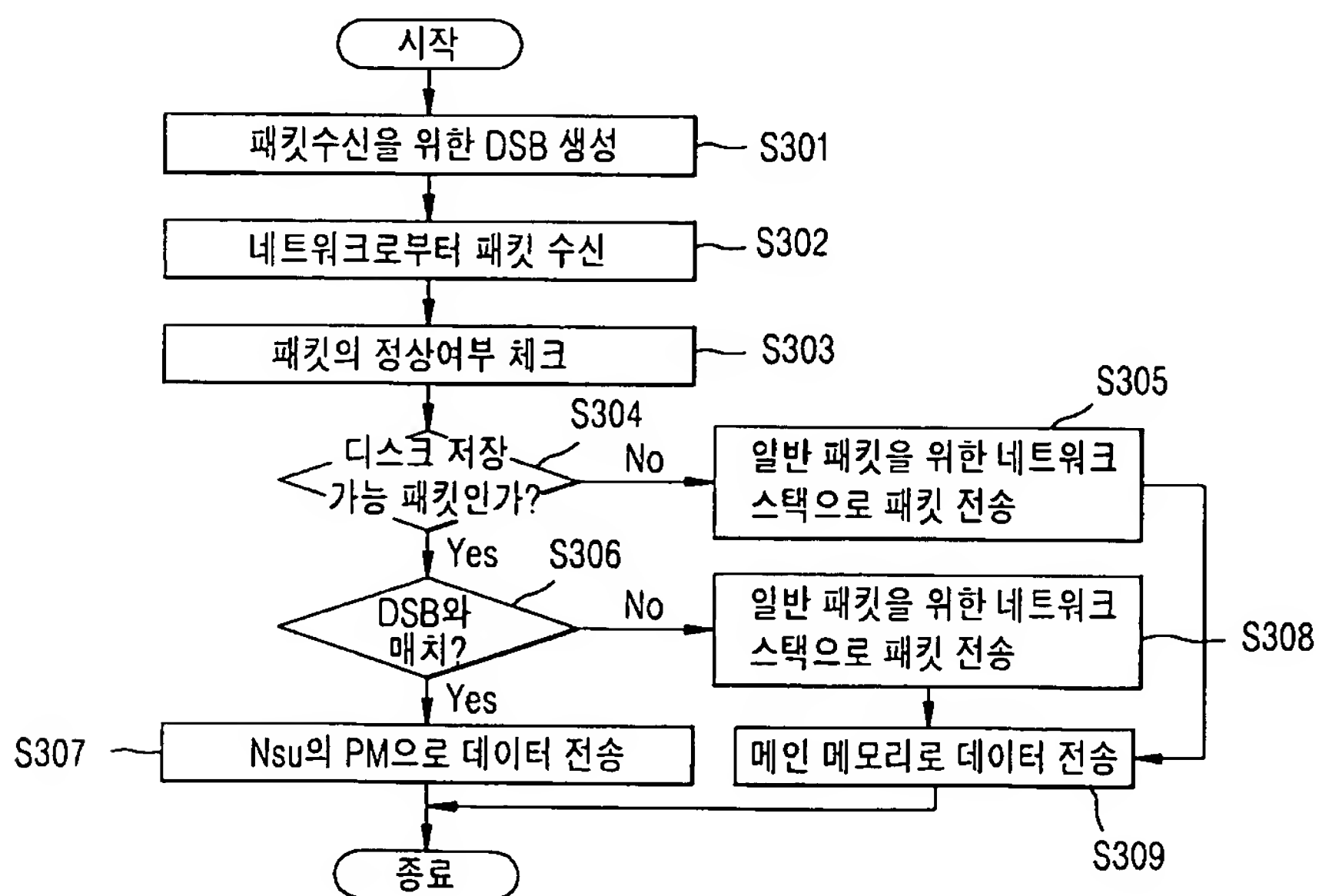
【도 1】



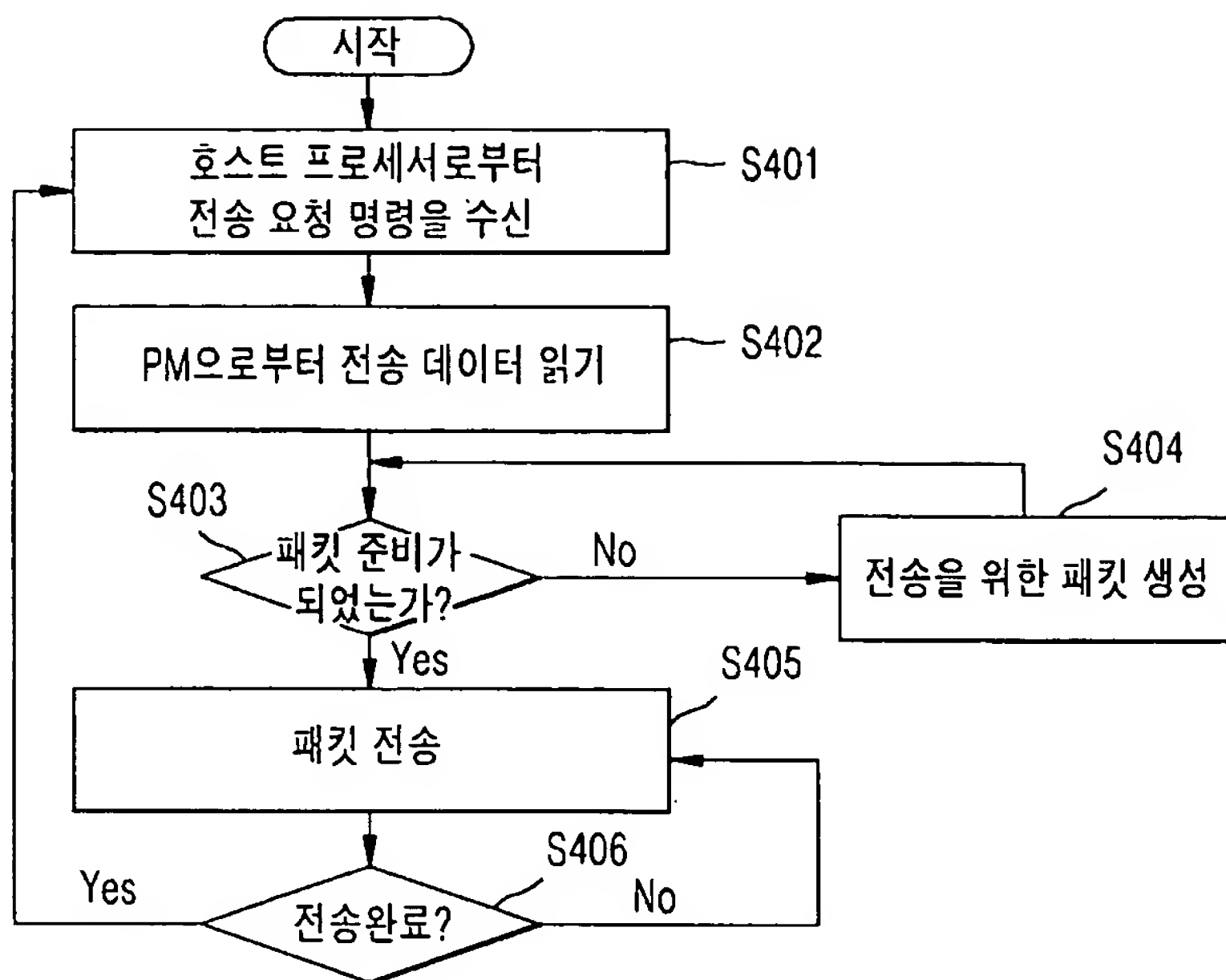
【도 2】



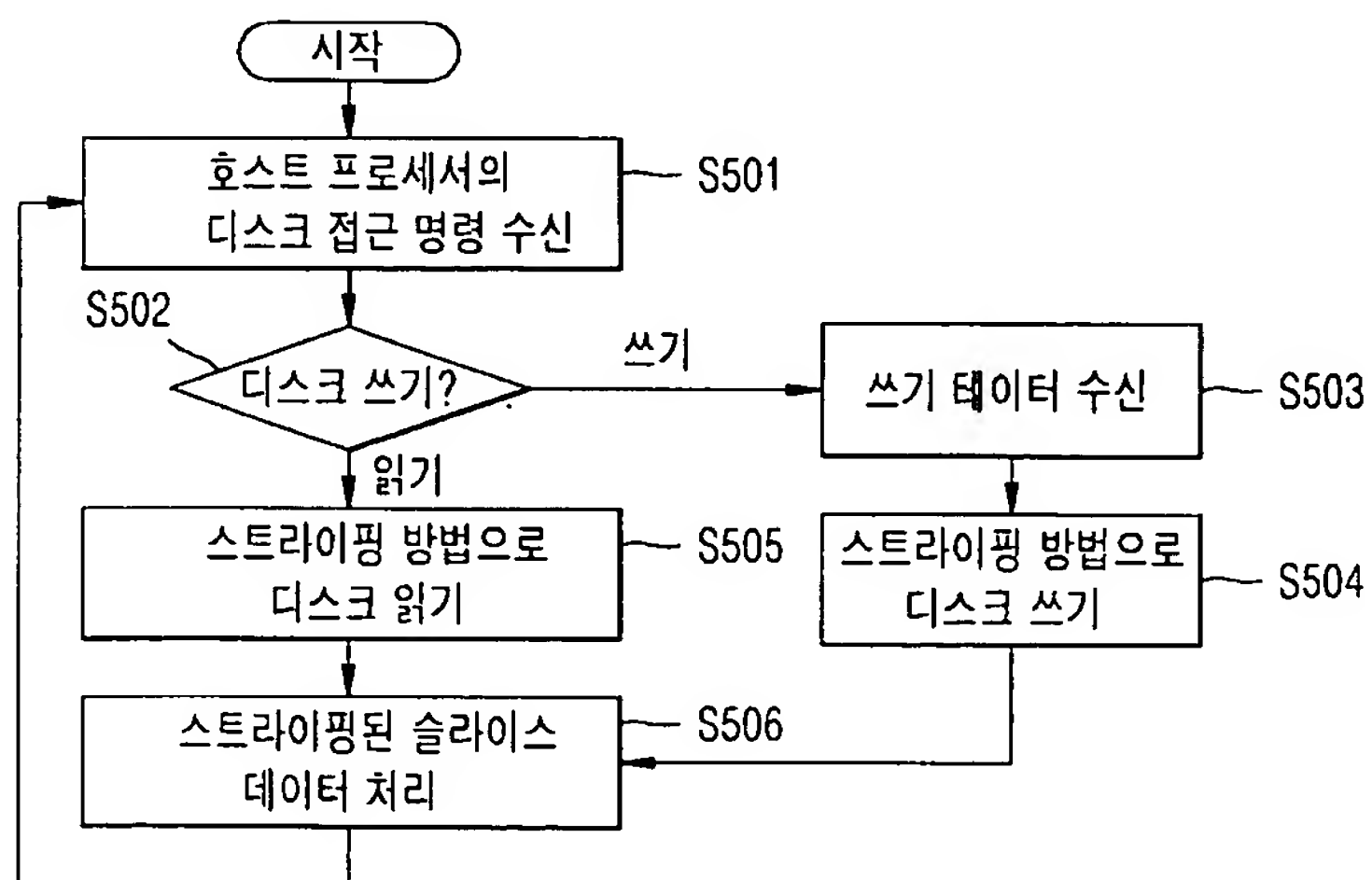
【도 3】



【도 4】



【도 5】



【도 6】

